

University of Groningen

Children's first and second-order false-belief reasoning in a verbal and a low-verbal task

Hollebrandse, Bart; van Hout, Angeliek; Hendriks, Petra

Published in:
Synthese

DOI:
[10.1007/s11229-012-0169-9](https://doi.org/10.1007/s11229-012-0169-9)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Final author's version (accepted by publisher, after peer review)

Publication date:
2014

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Hollebrandse, B., van Hout, A., & Hendriks, P. (2014). Children's first and second-order false-belief reasoning in a verbal and a low-verbal task. *Synthese*, 191(3), 321-333. <https://doi.org/10.1007/s11229-012-0169-9>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Children's first and second-order false-belief reasoning in a verbal and a low-verbal task

Bart Hollebrandse, Angeliek van Hout and Petra Hendriks

Abstract We can understand and act upon the beliefs of other people, even when these conflict with our own beliefs. Children's development of this ability, known as Theory of Mind, typically happens around age 4. Research using a looking-time paradigm, however, established that toddlers at the age of 15 months old pass a non-verbal false-belief task (Onishi and Baillargeon, 2005). This is well before the age at which children pass any of the verbal false-belief tasks. In this study we present a more complex case of false-belief reasoning with older children. We tested second-order reasoning, probing children's ability to handle the belief of one person about the belief of another person. We find just the opposite: 7-year-olds pass a verbal false-belief reasoning task, but fail on an equally complex low-verbal task. This finding suggests that language supports explicit reasoning about beliefs, perhaps by facilitating the cognitive system to keep track of beliefs attributed by people to other people.

Keywords Language · Theory of Mind · Second-order False Belief

1. Introduction

We understand and act upon the beliefs of other people, even when these are in conflict with our own beliefs. Children's development of this ability, known as Theory of Mind, has been extensively studied over the past twenty-five years, starting with the seminal study of Wimmer and Perner (1983). Theory of Mind (ToM) development involves various aspects of reasoning about others, including social awareness, joint attention, and anticipation of other people's behavior. Reasoning about false beliefs—the ability to handle the contrast between true and false beliefs—seems to develop rather late. Typically it is not until the age of four that children understand that, for instance, *John thinks that it is raining outside* contains a belief about the weather attributed to John, and are aware that John may be incorrect in his belief, in which case they must attribute a false belief to another person (Astington, 1993; Wellman and Bartsch, 1988). Our study involves more complex false-belief reasoning, adding another belief layer, as expressed in *Tom believes that John thinks that it is raining outside*.

Why does false-belief reasoning develop so late? Several studies have established a link with specific aspects of language development, suggesting a dependence between the two. Syntactic embedding is typically acquired around age 3 or 4, and precedes false-belief reasoning (de Villiers, 2005). De Villiers argues that the acquisition of the syntax of linguistic embedding, with verbs like *say* (e.g., *Mom said that it was rain-*

ing), is a prerequisite for developing the cognitive representations required for false beliefs. She bases her view on the observation that syntactic embedding is typically acquired around age 3 or 4, and precedes false-belief reasoning. Another aspect of language that correlates with the development of a Theory of Mind is the use of mental state verbs (Ruffman, Slade and Crowe, 2002). Furthermore, in the pragmatic domain the degree to which the child integrates his or her speech acts with previous speech acts in a conversation is correlated with the development of false belief reasoning (Dunn and Brophy, 2005). Theory of Mind thus develops hand in hand with various aspects of language and develops relatively late (see Milligan, Astington and Dack (2007) for an extensive overview).

But is false-belief reasoning really a late development? Tested on a non-verbal version of the so-called Sally-Ann task, children as young as 15 months pass a false-belief task (Onishi and Baillargeon, 2005). The researchers showed toddlers movies in which a toy is hidden in one location while an actor is watching; then the actor's view is blocked by a screen, and the toy is hidden in another location. When the screen opens again and the actor is about to reclaim the toy, the toddlers looked longer at the location where the toy was initially hidden (i.e., where the actor thinks the toy is), than at the location where the toy was hidden now. The children's looks reveal their expectation of the actor's behavior on the basis of that person's belief about the hiding place of the toy; this location is different from their own knowledge of the toy's loca-

tion. The children thus seem to track the actor's false belief about the location of a toy vis-à-vis their own, true beliefs (but see Perner and Ruffman, 2005).

Although one may doubt whether these tasks, which measure expectation, test the actual reasoning involved in considering false beliefs (for discussion, see Apperly and Butterfill, 2009; de Bruin and Newen, 2011), the Onishi and Baillargeon (2005) findings undoubtedly show that 15-month-olds effectively represent false beliefs. Apparently, toddlers develop implicit knowledge of false-belief attribution well before they can verbalize that knowledge explicitly, and before they pass any of the verbal false-belief tasks. Furthermore, these children pass this ToM test before they have acquired any complex syntax, thereby refuting the basis of De Villiers' (2005) hypothesis. Young children thus have some cognitive representation of false beliefs that does not rely on language.

In this study we present a more complex case of false-belief reasoning with older children. The case under investigation involves two layers of belief representations: the ability to understand one person's belief (first layer) about a belief attributed to another person (second layer), as in *Tom believes that John thinks that it is raining outside*, where Tom entertains the belief that John has a certain thought about the weather. Perner and Wimmer (1985) claim that this more complex ToM development is not mastered until the age of 7 or 8 (see also Sullivan, Zaitchik and Tager-Flusberg, 1994). Probing first-order and second-order ToM reasoning in 6 to 9-year-olds, we tested complex false-belief reasoning with a verbal task (questions after stories) and

an equally complex low-verbal task (questions after silent movie clips). We want to find out if, for such complex false-belief tasks, language supports the development of the cognitive representations of reasoning required to perform these tasks. We thus aim to test the hypothesis that language supports the development of higher-order false-belief reasoning. This hypothesis is more general than De Villiers' claim that the syntax for linguistic embedding is a prerequisite for false-belief reasoning and is compatible with Roeper's (2007) view of "language as a blackboard for thought".

3. Method

We used two tasks to test false-belief (FB) reasoning at first-order and second-order levels, the designs and materials of which were taken from the study of Hollebrandse, Hobbs, De Villiers and Roeper (2008) with English learners. The essence in both tasks is that the participants form beliefs about the situations that are different from the protagonists in the stories and video clips (first order); moreover, they know the beliefs of one protagonist about the beliefs of another protagonist, and how these are different (second order). The two tasks differ as to how the evidence for the beliefs was presented. In the verbal task, participants were told a story which provided the necessary clues for FB reasoning. In the low-verbal task, participants watched silent movies with one or two actors. The experimenter occasionally pointed out some features in the movies, but, crucially, there were no language clues about beliefs. Instead,

the clues for the beliefs of the different actors had to be deduced from the happenings in the visual context.

3.1 Participants

Forty-three Dutch children were tested, divided over two age groups: twenty-one 6 and 7-year-olds (mean age = 6;9, range = 6;2 – 7;3) and twenty-two 8 and 9-year-olds (mean age = 8;10, range = 8;2 – 9;11). We also tested a control group of seventeen adults. All subjects participated in both tasks. The data was collected in two sessions. The order in which the tasks were conducted was balanced across participants.

3.2 Verbal false-belief task

In the verbal task an elaborate story was told in which the beliefs of various people in the story were manipulated. The stories were accompanied by four pictures, which were presented one by one and served as a memory aid (see Appendix 1 for an illustration). The stories were modeled after Wimmer and Perner's (1985) "ice cream truck story". In contrast to their stories, we made sure that the beliefs of the two main protagonists in the story did not overlap, both at first-order and second-order level: each protagonist had his or her own, distinct belief which was different from that of the other protagonist, as well as from the belief of the participants.

All the stories have the same set up. Protagonist 1 and 2 initially share the same belief. In the sample story in Appendix 1, both main characters (Sam and Maria) initially think that there are chocolate-chip cookies at the bake sale of the church. Then pro-

tagonist 1's belief changes without protagonist 2 knowing about it (Sam's mom tells Sam that they are selling pumpkin pie). Next, protagonist 2 learns that the reality is different, without protagonist 1 knowing about this (Maria finds out that there are only brownies left). At this point protagonist 1 has a first-order belief which differs from his initial belief and also from the reality (Sam now thinks they're selling pumpkin pie, not chocolate-chip cookies; he doesn't know that in reality they're selling brownies). Protagonist 2 knows the reality, which is different from her second-order belief about protagonist 1 (Maria knows they're selling brownies, but thinks that Sam still thinks that they sell chocolate-chip cookies).

We did not use any second-order, syntactic embedding constructions of the type *Maria thinks that Sam thinks they are selling cookies at the bake sale* in the story. Instead we elicited a second-order answer by asking a "double" first-order question. In the bake-sale story the mailman asks Maria a first-order question *What does Sam think they are selling at the bake sale?* The experimenter then asks the participant what Maria will say to the mailman (see also Sullivan, Zaitchik and Tager-Flusberg, 1994). The child thus did not need to process any second-order embedding structures in language, but the task still involved second-order reasoning.

There were eight stories of this format, each containing one second-order question and two first-order ones. The initial first-order question was asked in the middle of the story; the same question was asked once again at the end of the story. The purpose of asking the same question twice was to check whether children had difficulties with

the length and complexity of the story. The repeated first-order question thus effectively served as a control of how well participants were able to keep track of the different beliefs despite the length and complexity of the story.

3.3 Low-verbal false-belief task

For the low-verbal task, participants also had to keep track of the different beliefs of different protagonists in the same situation. Whereas the former task was fully verbal, this one limited the use of language as much as possible. The experimenter only drew attention to the contents of the box (or the location of the object) and pointed out the screens going up and down. This was done without using any propositional-attitude verbs (such as *think* or *believe*), and without referring to thoughts or beliefs in any other way.

Participants saw short movies. In the movies one or two observers watched a changing scene from behind a window. In half of the movies the contents of a box were changed three times (following the Unknown-Change-of-Content scenario of Wimmer and Perner's (1983) *Smarties task*), and in the other half an object was moved between three different locations (following the Unknown-Change-of-Location scenario of Wimmer and Perner's (1983) *Sally-Ann task*). At certain moments during the changes, a screen was lowered down the window so that the observers could not see the scene, thus missing a crucial part of the changes. Then the screen went up again. The screen lowering happened at different moments for both observers. In this way

we created a situation where one observer has a different belief of the contents of a box (or the location of the object) than the other observer. The participant has an overall view and could know who believes what (see Appendix 2 for an illustration).

The task was presented as a game in two parts. In the first part the participant herself was the player of the game and had to keep track of what the single observer in the movie knows about the changes of the contents of a box (or the different locations in the change-of-location variant of this task). At the end of the series of changes, the experimenter asks the participant what the observer thinks is in the box. These were the first-order trials (see Appendix 2a).

In the second part participants were told that it was the same game, but there was an additional observer in the movie (i.e., the man in the window on the right in Appendix 2b) who was the player now. The participant had to keep track of what this observer knew about the contents of the box, and what he knew about the first observer's beliefs (the woman). The series of scene changes in the second-order movies was the same as in the first-order movies (the contents of a box changes three times, or a toy moves between three locations). This task involves second-order reasoning: by lowering the screen of the two observers at different moments, we created a second-order false belief for the man about the woman. For example, the man would incorrectly believe that the woman thinks there is an apple in the box, whereas she actually believes (based on what she saw) that there is a small basket in the box; in reality, however, there is a turtle in the box.

Four movies tested first-order FB reasoning and four others tested second-order FB reasoning. For the younger children, a first-order question was added at the end of the second-order trials to check whether they were able to follow the complicated series of events.¹

3.4 Similarities and differences between the two tasks

In both tasks participants had to reason with first-order and second-order false beliefs, as the beliefs of the two protagonists differed from each other as well as from the participants themselves. Furthermore, participants had to keep track of two first-order beliefs and one second-order belief: in the story task, they had to keep track of Sam's and Maria's first-order beliefs as well as Maria's second-order belief about Sam's first-order belief. In the movie task, they had to keep track of the first-order beliefs of the protagonist in the right window and the protagonist in the left window, and the second-order belief of the protagonist in the right window about the belief of the other protagonist in the left window.

The tasks differed in that the clues about who believes what about whom were either presented in a story, or had to be deduced from the happenings in the movies, hence our labels "verbal" versus "low-verbal" task. We call the movie task "low-verbal" because it was not totally non-verbal. The experimenter draws verbal attention to the changes of the contents of the box or the location of the object; moreover, the prompt at the end of each movie was also verbal ("What does she think is in the

box?”, “What will he say?”). Conversely, the verbal task was supported with pictures. The two tasks also differed in the number of protagonists; the low-verbal task has only two protagonists (the woman and the man), whereas the verbal task has two main characters (Sam and Maria) and two additional characters (the mom and the mailman). Having more protagonists adds to the complexity of the mental representations involved. Potentially it makes the verbal task more demanding. However, as we will see in the next section, this is not reflected in the results. Children are more accurate on the verbal task than the low-verbal task.

4. Results

The adults performed nearly at ceiling at all test questions, with 96% correct responses on the second-order question in the verbal task and 91% in the low-verbal task. Figures 1 and 2 present the children’s scores on the verbal and low-verbal tasks. For both, the results show a sharp difference between first-order and second-order questions. Moreover, for the second-order items, children performed better in the verbal task than in the low-verbal task.

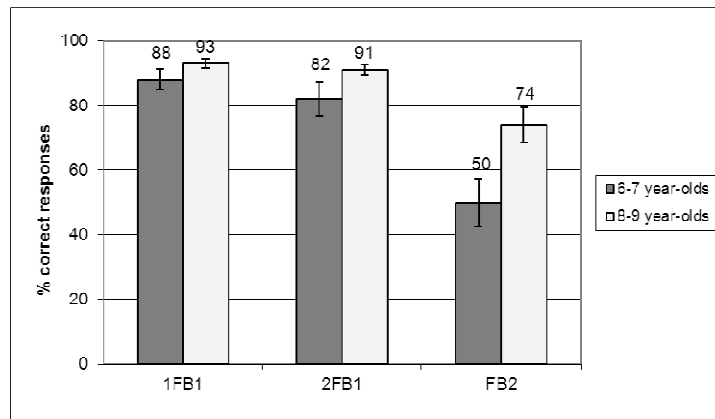


Figure 1: Verbal false-belief task: Percentage of correct responses for both age groups (error bars show standard errors) on 1FB1 (first-order FB question at first time), 2FB1 (first-order FB question at second time) and FB2 (second-order FB question).

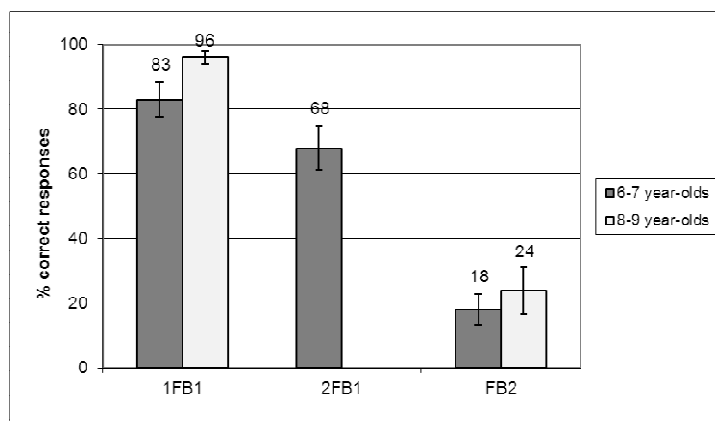


Figure 2: Low-verbal false-belief task: Percentage of correct responses for both age groups (error bars show standard errors) on 1FB1 (first-order FB question in first-order trials), 2FB1 (first-order FB question in the second-order trial) and FB2 (se-

cond-order FB question in second-order trials). Note that the older children did not receive the 2FB1 question.

Mixed ANOVAs were performed with Presentation Mode (verbal – low-verbal) and False Belief Level (first order – second order) as within-participants factors, and Age (younger children – older children) as the between-participants factor. There were main effects for Presentation Mode, $F(2,43) = 51.4$; $p < 0.001$; $\eta_p^2 = 0.556$) and False Belief Level, $F(2,43) = 160.7$; $p < 0.001$; $\eta_p^2 = 0.797$), and a significant interaction between the two ($F(2,43) = 37.4$; $p < 0.001$; $\eta_p^2 = 0.477$). The children performed better on the verbal than on the low-verbal task, and they also performed better on first-order questions than on second-order questions. Their performance on first-order items was higher than on second-order items in both tasks, but performance on second-order questions was particularly low in the low-verbal task as compared to the verbal task.

Taking a closer look at these effects, we performed paired-sample t-tests. There were significant differences between the verbal and the low-verbal second-order responses, both for the younger group ($t(20) = 4.1$; $p = 0.003$ (Bonferroni corrected) and the older group ($t(20) = 6.5$; $p < 0.001$). There were no significant differences between any of the first-order false-belief responses. This general lack of effect at first-order level indicates that participants had no problem at this level of reasoning. The differences between the tasks lie instead at the second-order level of reasoning.

Furthermore, age was also significant ($F(2,43) = 7.9$; $p = 0.008$; $\eta_p^2 = 0.161$). This main effect could not be pinned down to any of the specific differences.

5. Discussion

Onishi and Baillargeon (2005) show that implicit non-verbal FB reasoning is accomplished at a very young age, which suggests that language is not required for the implicit representation of first-order false beliefs. Explicit verbal FB reasoning is acquired around the age of four, with the help of language. In this study, we investigated second-order reasoning about false beliefs with a verbal and a low-verbal task, comparing it to first-order FB reasoning. The goal was to find out if language plays a crucial role in higher-order false-belief reasoning.

Children's success on the first-order false-belief items in the two tasks indicates that they were able to keep track of the different beliefs despite the complexity of the set-up with several protagonists, each with their own changing beliefs. Task complexity may have had some effect, however, as children's performance on the first-order question in the low-verbal task was found to be slightly lower in the second-order condition with two observers than in the first-order condition with only one observer. Still, children's performance was well above chance: at 33.3% chance $t(20) = 5.2$; $p < 0.001$ (and at 50% $t(20) = 2.7$; $p = 0.015$).²

Children's near-ceiling performance on the first-order items contrasts with their much poorer performance on the second-order items. The verbal second-order task

also turned out to be easier than the corresponding low-verbal task. Thus we have established a crucial effect of mode of presentation in the representation of complex beliefs at second-order level, which suggests a role for language in the construction of these representations. Why is the verbal task easier for children? And what does this suggest about the relation between language development and false-belief reasoning?

A first possibility is that children's different performance on the two tasks is a task effect unrelated to their capacity for false-belief reasoning. Retrieval of the relevant belief representations from memory in order to respond to the test question may be aided by the manner in which these representations have been processed and stored during the task. In the verbal task the story was presented verbally. Moreover, there were probe questions at various moments throughout the story which were designed to prompt participants to explicitly remember (and verbalize) their knowledge about the ongoing shifts in false beliefs. In the low-verbal task, on the other hand, the story was presented purely visually and there were no probe questions. Whereas in both tasks, participants had to infer the changing belief states of the protagonists, the probe questions in the verbal task forced them to explicitly track those beliefs. It is thus conceivable that the explicit nature of the verbal task makes it an easier task for children. A way to investigate this possibility further would be to design an alternative version of our low-verbal task, adding probe questions as the movies unfold, asking participants explicitly to remember who believes what and verbalize this knowledge. However, the conclusion that an explicit verbal reasoning task is essentially easier than an

implicit one, is at odds with the observed difference between young children's good performance on low-verbal, implicit false-belief tasks versus their difficulty with explicit false-belief tasks (see Section 1).

An alternative explanation related to task effects is that the mismatch between mode of presentation and mode of response makes the low-verbal task more difficult for children. In both tasks a verbal response was required in response to the test question ("What does she think is in the box?"). In the low-verbal task, therefore, participants are forced to switch modalities, which may increase processing costs (Dantzig, Pecher, Zeelenberg and Barsalou, 2008). If this explanation is on the right track, we predict that children will perform better when the task is entirely non-verbal and is measured by looking behavior, as in an eye-tracking study. The methodology of eye-tracking has the advantage that the experimental measure is in the same mode as the stimulus presentation. In such a paradigm language is not an interfering factor. Using a change-of-location set-up, children's eye gaze to the three possible hiding locations of an object could thus be measured and compared.

If this mode-of-presentation explanation is correct, it is expected that in implicit higher-order reasoning tasks that do not rely on language either in their presentation or for their response, children may not experience the same difficulty as in our low-verbal task. However, non-verbal higher-order reasoning tasks such as required in strategic games are not easy either. Hedden and Zhang (2002) and Meijering, van Maanen, van Rijn and Verbrugge (2010) argue that implicit second-order reasoning

does not come readily for adults. Furthermore, Flobbe, Verbrugge, Hendriks and Krämer (2008) find that children perform better at a verbal second-order false belief task than at a non-verbal second-order strategic game. These studies show that participants do not perform at ceiling in non-verbal modes of presentation, such as with strategic games. This suggests that modes of presentation other than language are not necessarily easier and that language perhaps is a supporting factor in second-order false-belief reasoning.

This brings us to a third possibility, which is that children perform better on the verbal task because they are helped in their false-belief reasoning by language. On a syntactic view, the ideas of De Villiers (2005) about the role of linguistic embedding in acquiring explicit first-order false-belief representations may be extended to second-order false-belief reasoning. Second-order false-belief reasoning requires recursive embedding of beliefs in a way similar to how language recursively structures syntactic embedding. The recursive nature of the linguistic representations involved in syntactic embedding may therefore provide a scaffolding for constructing recursive belief structures which are required to perform the recursive step in a second-order false-belief reasoning task (Hollebrandse and Roeper, submitted). On the other hand, the results of our study are also compatible with the possibility that semantic or pragmatic rather than syntactic aspects of language help children in their false-belief reasoning.

We conclude that a verbal second-order FB task is easier for children than a corresponding low-verbal FB task. This suggests that language supports explicit reasoning about beliefs, perhaps by facilitating the cognitive system to keep track of beliefs attributed by people to other people.

Acknowledgements

The tests used in this study have been developed by the first author in collaboration with Tom Roeper, Jill and Peter de Villiers and Kate Hobbs. We thank the children and teachers of the Annie M.G. Schmidt School and the Joseph Haydn School in Groningen for their hospitality. This paper has benefitted from the comments of three anonymous reviewers and the audience of the TARK 2011 workshop (University of Groningen), and the audience at the 2012 workshop Language Development in Childhood and Adolescence (University of Leiden). We also thank two anonymous reviewers of *Synthese* for their detailed and constructive comments. We gratefully acknowledge ESF (grant no. 028395, PI M. Krifka and U. Sauerland) and NWO (grant no. 277-70-005, PI P. Hendriks).

References

Apperly, I. A. and Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *American Psychological Association*, 116, 953-970 (2009)

Astington, J. W. (1993). *The child's discovery of the mind*. (Cambridge: Harvard University Press)

Bruin, L. de & Newen A. (2011). The Developmental Paradox of False Belief Understanding: A Dual-system Solution. (In J. van Eijck & R. Verbrugge. (Eds.), *Reasoning about Other Minds: Logical and Cognitive Perspectives*, CEUR Workshop Proceedings (pp. 32-49). <http://ceur-ws.org/Vol-751/>

Dantzig, S., Pecher, D., Zeelenberg, R. & Barsalou, L. (2008). Perceptual processing affects conceptual processing. *Cognitive Science*, 32, 579-590

De Villiers, J. G. (2005). Can language acquisition give children a point of view? (In J. Astington, & J. Baird (Eds.), *Why Language Matters for Theory of Mind* (pp. 186-219). Oxford: Oxford Press)

Dunn, J. & Brophy, M. (2005). Communication, relationships, and individual differences in children's understanding of mind. (In J. Astington, & J. Baird (Eds.), *Why Language Matters for Theory of Mind* (pp. 186-219). Oxford: Oxford Press)

Flobbe, F., Verbrugge, R., Hendriks, P. & Krämer, I. (2008). Children's application of Theory of Mind in reasoning and language. *Journal of Logic, Language and Information*, 17:4, 417-442

Hedden, T. & Zhang, J. (2002). What do you think I think you think?: Strategic reasoning in matrix games. *Cognition*, 85, 1-36

Hollebrandse, B & Roeper, T. Recursion and propositional exclusivity. (submitted)

Hollebrandse, B., Hobbs, K., De Villiers, J. G. & Roeper, T. (2008). Second order embedding and second order false belief. (In A. Gavarro, & M. J. Freitas (Eds.), *Language Acquisition and Development, Proceedings of GALA 2007* (pp. 270-280). Cambridge: Cambridge Scholar Press.)

Meijering, B., Maanen, L., van, Rijn, H. van & Verbrugge, R. (2010). The facilitative effect of context on second-order social reasoning. (In R. Catrambone, & R. Ohlsson (Eds.), *Proceeding of the 32nd Annual conference of the Cognitive Science Society*, (pp. 1423-1428). Austin: Cognitive Science Society.)

Milligan, K., Astington, J. W. & Dack, L. A. (2007). Language and theory of mind: meta-analysis of the relation between language ability and false-belief understanding. *Child Development*, 78, 622-646

Onishi, K. H. & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308, 255-258

Perner, J. & Wimmer, H. (1985). "John thinks that Mary thinks that" Attribution of second order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology*, 39, 437-471

Perner, J. & Ruffman, T. (2005). Infants' insight into the mind: how deep? *Science*, 308, 214-216

Roeper, T. (2007). *The prism of grammar. How child language illuminates humanism.* (Cambridge: M.I.T. press)

Sullivan, K., Zaitchik, D. & Tager-Flusberg, H. (1994). Preschoolers can attribute second order beliefs. *Developmental Psychology*, 30, 395-402

Wellman, H. M., & Bartsch, K. (1988). Young children's reasoning about beliefs. *Cognition*, 30, 239-277

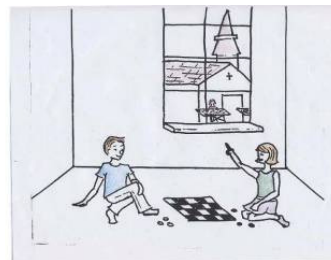
Ruffman, T., Slade, L. & Crowe E. (2002). The relation between children's and mothers' mental state language and theory-of-mind understanding. *Child Development*, 73, 734-751

Wimmer, H. & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13, 103-128

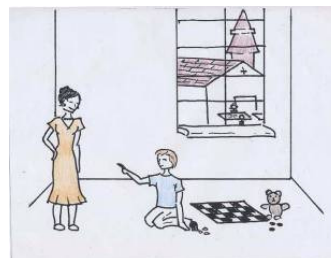
Appendix 1: Illustration of verbal false belief task (the *Bake Sale Story*)

Abbreviations: Q1FB1 = initial first-order false-belief question; Q2FB1 = repeated first-order false-belief question; QFB2 = second-order false-belief question.

Sam and Maria are playing together. They look outside and see that the church is having a bake sale. Maria tells Sam: “I am going to buy chocolate chip cookies for us there,” and she walks away.

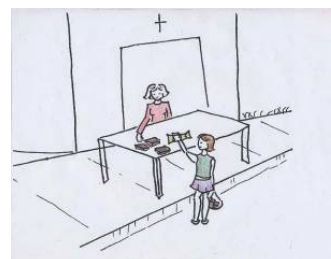


Mom comes home and she tells Sam that she just drove past the bake sale. “Are they selling chocolate chip cookies?” Sam asks. “No,” mum says, “they are only selling pumpkin pie.” “Maria will now probably get pumpkin pie at the bake sale,” Sam says.



Probe 1: Does Maria know they are selling pumpkin pie at the bake sale?

Maria has arrived at the bake sale. “I would like to buy chocolate chip cookies,” she says. “All we have left are brownies,” says the lady



behind the stall. Since Maria also likes brownies, she decides to get some brownies.

Probe 2: Does Sam know that Maria bought some brownies?

Q1FB1: What does Sam think they are selling at the bake sale? Why does he think that?

On her way back, Maria meets the mailman.

She tells the mailman: “I have just bought some brownies. I am going to share them with my brother Sam. It is a surprise”. “That is nice of you,” says the mailman. Then he asks Maria: “Does Sam know what you bought him?”



Ignorance: What does Maria tell the mailman?

Then the mailman asks: “What does Sam think they are selling at the bake sale?”

QFB2: What does Maria tell the mailman? Why does she say that?

Q2FB1: What does Sam think they are selling at the bake sale?

Why does he think that?

Appendix 2: Illustration of non-verbal false-belief task (Unknown-Change-of-Content set-up)

Abbreviations: Q1FB1 = first-order false-belief question in the first-order task; Q2FB1 = first-order false-belief question in the second-order task; QFB2 = second-order false-belief question.

a. Non-verbal first-order task



First order trial

Q1FB1: Remember you are the player. Now, what does she think is in the box?
(experimenter points to the woman)

b. Non-verbal second-order task



Second order trial

QFB2: Remember, first you were the player, but now he (experimenter points to the man in the right window) is the player and we are going to ask him the same question as we asked you: “What does she think is in the box?” What will he answer?

Q2FB1: What does she herself think is in the box?

Examples of non-verbal movies can be found at:

<http://www.let.rug.nl/hollebr/FB-movies/1stOrderNonVerbal.wmv>

<http://www.let.rug.nl/hollebr/FB-movies/2ndOrderNonVerbal.wmv>

Notes

¹ The second first-order question was not asked with the older children as they were tested before the younger children. At that time, we feared that adding another question would make the task too demanding. However, as the older children turned out to be quite successful with the second first-order question in the verbal task, our fears appeared to be unwarranted. We then decided to add the second first-order question to the non-verbal task for the younger children as well, providing us with a measure of complexity.

² Chance was calculated by determining whether actual scores significantly differed from 33.3% or 50%.